面向数字微流控生物芯片的差分隐私方案

陈 潇[1,2], 董 晨[1,3]

[1] 福州大学 计算机与大数据学院/软件学院 福州 中国 350116
[2] 网络系统信息安全福建省高校重点实验室 福州 中国 350116
[3] 福建省网络计算与智能信息处理重点实验室 福州 中国 350116

摘要 近年来, 基于数字微流控生物芯片(Digital Microfluidic Biochip, DMFB)的分子诊断技术成为热点研究方向。与传统分子诊断技术相比, 数字微流控生物芯片具有精准控制离散液滴、执行生化协议等优势。然而, 作为网络物理系统的组成, 生物芯片潜在的隐私安全问题日益突出, 比如通信信道的窃听攻击, 生化协议的篡改攻击, 物理结构保护的版权攻击等。差分隐私作为传统数据隐私保护的常用技术可以融入生物芯片应用以保护用户隐私安全。然而, 对隐私安全、生物芯片应用以及生物芯片安全三种技术的交叉研究较为少见。调研分析生物芯片应用的实现机制和威胁模型, 包括生化协议、网络物理系统及增强隐私保护的 DMFB 用户数据安全平台, 首先在 DMFB 用户数据平台上描述了拉普拉斯机制、高斯机制和随机响应机制的应用场景和保护方案, 其次基于用户层级敏感度、路由权重集合和路由交叉点参数集合这三个策略提出参数安全发布算法, 最后创建防篡改概率作为安全性指标, 同时建立置信分数、校准度和累计误差率衡量数据可用性。仿真实验结果表明整体方案的隐私安全性可达 98%, 数据可用性平均可达 93.3%, 算法性能试验表明方案最佳的隐私预算为 0.4, 此外, 对比同类算法, 所提方案平均提高了 12.09% 隐私安全性和 7.02%的数据可用性, 因此该方案能够为 DMFB 执行生化协议安全有效的用户数据平台。

关键词 数字微流控生物芯片; 生化协议; 数据安全; 差分隐私
中图法分类号 TP309.2 DOI 号 10.19363/J.cnki.cn10-1380/tn.2024.11.04

Differential Privacy Scheme for Digital Microfluidic Biochips

CHEN Xiao[1,2], DONG Chen[1,3]

[1]College of Computer and Data Science/College of Software, Fuzhou University, Fuzhou 350116, China
[2] Key Lab of Information Security of Network Systems, Fujian Province, Fuzhou 350116, China
[3] Fujian Key Laboratory of Network Computing and Intelligent Information Processing, Fuzhou University, Fuzhou 350116, China

Abstract Digital Microfluidic Biochip (DMFB) -based molecular diagnostic techniques have recently become hot topics. Compared with traditional molecular diagnostic techniques, digital microfluidic biochips have advantages in precise control of discrete droplets and execution of biochemical protocols. However, as components of networked cyber-physical systems, potential privacy and security issues of biochips are increasingly prominent, for instance, eavesdropping attacks on communication channels, tampering attacks on biochemical protocols, and copyright attacks on physical structure protection. Differential Privacy (DP), a de facto standard for achieving privacy, is trying to incorporate DMFB applications to protect user privacy. However, as the intersection of privacy-preserving technology, DMFB applications, and DMFB security, comprehensive research on this area is relatively rare. Investigating and analyzing the implementation mechanisms and threat models of biochip applications, including biochemical protocols, cyber-physical systems, and Enhanced privacy protection DMFB's user data security platform, this paper proposes the application scenarios and protection schemes of differential privacy techniques on DMFB user data platform. Firstly, the application scenarios and protection schemes of Laplace mechanism, Gaussian mechanism, and random response mechanism were described on the DMFB user data platform. Secondly, parameter security publishing algorithms were proposed based on three strategies: user level sensitivity, routing weight set, and routing intersection parameter set. Finally, tamper proof probability was created as a security indicator, while confidence scores, calibration, and cumulative error rate were established to measure data availability. The simulation experiment results show that the overall privacy security of the scheme can reach 100%, and the average data availability can reach 93.3%. The algorithm performance test shows that the optimal privacy budget range of the scheme is 0.4. In addition, compared with similar algorithms, the proposed scheme improves privacy security by 12.09% on average, and data availability by 7.02%. Therefore, this scheme can be a secure and effective user data platform for DMFB to execute biochemical protocols.

通讯作者: 董晨, 博士, 讲师, Email: dongchen@fzu.edu.cn
本课题得到福建省高校数字经济学科联盟建设经费, 福建省自然科学基金(No. 2020J01500)资助。
收稿日期: 2023-03-03; 修改日期: 2023-06-17; 定稿日期: 2024-09-05

Key words digital microfluidic biochip; biochemical protocol; data security; differential privacy

1 引言

数字微流控生物芯片(Digital Microfluidic Biochip, DMFB)是一种被称为“芯片上实验室”的新兴技术, 具有样本量小、集成度高等优点, 有效降低样本和试剂消耗率并实现自动化检验分析[1], 实现临床护理的小型化分析诊断、DNA 测序和环境监测[2]。医学诊断是微流体研究的一个重要应用领域, DMFB 生物芯片具备疾病快速诊断检测能力, 有利于提高资源匮乏地区的医疗服务质量[3]。DMFB 网络物理系统发送和接收用户数据, 并为决策执行数据挖掘, 存在着隐私安全漏洞[1]。2020 年 12 月, 国家发展和改革委员会, 工业和信息化部、中国网络空间局和能源局联合发布了《关于加快建设国家一体化大数据中心协同创新体系的指导意见》, 在大数据开发系统中, 数据隐私保护发挥着重要作用。

1.1 隐私保护背景

全球生物芯片市场规模预计将从 2018 年的 57 亿美元增长到2025年的123亿美元。Babies' SEEKER 新生儿分析仪是一种基于 DMFB 的免疫分析平台, 2016 年获得 FDA 批准, 至 2019 年已运送共计 300 万份的测试样本, 筹集了 1300 万美元的资金。用于单细胞分析的 10x Genomics 将微流体和生物信息学相结合, 自 2012 年成立以来, 直到 2018 年它已经获得了 2.43 亿美元的资金。

随着生物芯片进入市场, 生物芯片逐渐成为医疗保健服务不可或缺的一部分, 生物芯片网络物理系统由硬件、软件和网络连接组成, 类似于当前的医疗设备, DMFB 存在篡改控制、拒绝服务和数据盗窃等安全漏洞导致大量医疗设备的召回, 将衍生出用户隐私隐患, 因此针对隐私保护的研究至关重要。

表 1 为 DMFB 数据库片段示意表, 以此为例描述 DMFB 数据收集任务场景以及可能存在的隐私威胁。在 DMFB 数据库中, 记录来自 DMFB 执行生化协议的实时数据和最终数据, 研究人员可以分析数据库中包含的模式规律。例如, 统计问题“数据库中编号为2116的病人有多少条检测记录满足属性P?”属性 P 可以是“特异性蛋白的含量不超过最大值”或者“血液中含氧量不小于最小值”。这些基于个人信息的查询, 会将个人隐私直接暴露出来。例如在片段数据中, 并没有直接显示用户的敏感信息, 但根据病人编号可以直接搜索到病人的基本信息; 以及具有背景知识的攻击者, 只需要掌握其中某些“检测类型”的结论, 就可以推断该用户的就医信息和检测结果, 这些医疗诊断的检验结果通常是重要的医疗健康数据, 这样潜在的隐私信息泄露会导致巨大社会危害性。

如今, 差分隐私已成为权威机构执行隐私保护的实际标准, 差分隐私技术在数学上有严格的定义且具有隐私的稳健性, 满足其定义的需求和算法也在增加[4]。随着 DMFB 在医疗诊断领域的研究越发深入, 对测定结果、校准曲线和控制信号等生化协议执行过程中的中间参数会进行更强大的收集和管理, 然而面对 DMFB 生物芯片安全漏洞和恶意程序的篡改攻击, 鲜有学者注意到差分隐私技术可以试图融入生物芯片应用场景以应对与 DMFB 生物芯片日俱增的数据安全威胁。因此本文应用并改进差分隐私保护技术中的实现机制, 对用户测定数据进行保护, 包括液滴测定结果、传感器校准范围和控制器反馈信号, 并提出基于层级敏感度的参数安全发布算法。综合实验表明 DMFB 差分隐私保护方案在液滴测定结果、传感器校准范围和控制器反馈信号上表现出了较高的安全性, 对比实验验证差分隐私方案保护生化协议参数安全的可用性。综上所述, 改进的差分隐私保护技术给 DMFB 生物芯片安全提供了一个可行方案。

表 1 DMFB 用户数据库

Table1 DMFB user database

DMFB 型号	病人编号	检测类型	片段协议	时间/Hrs	传感器	单位	液滴浓度	单位	最大值	最小值	反馈信号
D1100	2116	免疫测定	1	1200	1	mg/dl	200	mL	180	540	1
D1100	2116	免疫测定	2	1201	1	mg/dl	201	mL	180	540	1
D1100	2116	免疫测定	3	1202	2	mg/dl	202	mL	60	120	1
D1100	2116	免疫测定	4	1203	2	mg/dl	203	mL	120	180	1
D1100	2116	免疫测定	5	1204	2	mg/dl	204	mL	120	180	1
D1101	2117	免疫测定	1	1205	1	mg/dl	205	mL	60	120	1
D1101	2117	免疫测定	2	1206	1	mg/dl	206	mL	120	180	1

1.2 相关研究介绍

表 2 对防御类型和攻击来源、阶段、对象以及代表文献进行总结。目前为止, 面对安全威胁, 学术界的主要应对措施是提高DMFB系统可靠性[5-9]和对恶意程序的识别能力[9-13], 有效防止窃听、篡改等攻击方式。其中, DMFB 的可靠性研究包括设计验证方案[5]、数字水印方案[6]、芯片认证方案[7]和知识产权保护方案[8]以及 DMFB 的篡改攻击对象包括样品浓度[9]、测定参数[8]、混合-分裂操作[13]和校准曲线[9, 11]。

文献[5]使用LSB水印和AES高级加密标准对生物芯片设计进行认证和加密; 文献[6]使用数字签名对生化协议合成参数进行保护; 文献[8]改变PB-DMFB 的物理结构来对数字版权进行管理; 文献[7]基于生物芯片的物理不可克隆性, 生成电极固有变化的安全密钥; 文献[14]基于微流体多路复用器, 生成受微流体模式控制的安全密钥。

表 2 DMFB 生物芯片的安全威胁

Table 2 Security threats of DMFB biochips

防御类型	攻击来源	攻击阶段	攻击对象	代表文献
DMFB 生物芯片可靠性	不可信的通信信道	芯片设计	芯片布局设计信息	[5]
	第三方制造商和外包设计	生产制造	生化协议合成参数	[6]
		生产制造	芯片认证可信度	[7]
		生产制造	数字版权管理	[8]
	硬件木马、生产过剩和伪造攻击	终端用户	液滴和生化协议	[14]
恶意程序的防御能力	硬件木马带来的篡改攻击	标准化编码	驱动序列	[10]
	篡改攻击	芯片设计、生产制造	校正曲线	[11]
	不可信外围设备带来的窃听攻击	终端用户	自动逆向工程	[12]
	测定结果篡改攻击	终端用户	原始样品浓度、浓度校准曲线	[9]
	生化协议篡改攻击	终端用户	液滴的混合-分裂操作	[13]
	恶意液滴篡改攻击	终端用户	随机检查点	[15]

文献[10]基于汉明距离, 度量驱动序列的变化程度以应对恶意程序攻击; 文献[13]使用虚拟混合-分裂操作锁定生化协议以应对篡改攻击; 文献[12]提出基于混淆、伪装策略的自动逆向工程 BioChipWork; 文献[9]提出葡萄糖测定中的潜在攻击方案, 防止样品的原始浓度和校准曲线被篡改; 文献[11] 提出DMFB 体外免疫测定安全模型以应对参数、试剂和校准曲线的篡改攻击; 文献[15]利用基于随机检查点的入侵防御系统来提高对于恶意液滴的防御能力。

1.3 主要工作及创新点

在国内外差分隐私技术领域的研究基础上, 尝试应用差分隐私技术来研究数字微流控生物芯片的数据安全策略, 对当前DMFB生物芯片在分子诊断领域的潜在隐私安全问题[16]提出了可能的解决方案。该方案在液滴、传感器校准参数和控制器的反馈信号参数中分别嵌入拉普拉斯扰动、高斯机制扰动和随机响应扰动, 并利用三类参数建立层级敏感度, 以及对DMFB 的液滴路由和路由交叉点进行隐私预算分配, 从而建立了参数安全发布算法, 保证用户数据安全。

仿真实验结果表明整体方案的隐私安全性可达100%, 数据可用性平均可达 93.3%, 算法性能试验表明方案最佳的隐私预算范围为 0.4, 对比文献[6]和文献[17], 差分隐私机制提高了生化协议测定参数 12.09%的隐私安全性, 同时增加 7.02%的数据可用性。

2 相关技术概述

本章在DMFB生物芯片和DMFB网络物理系统的基础上建立了增加隐私保护的用户数据平台。

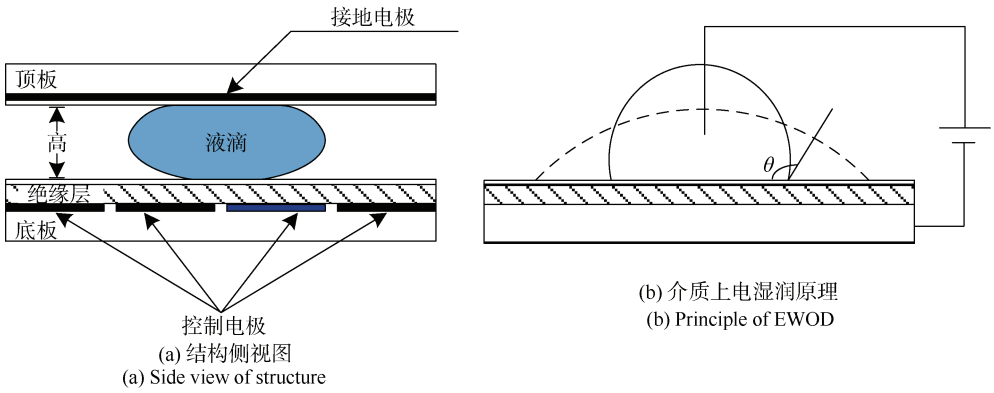
2.1 DMFB 生物芯片

图 1(a)是DMFB生物芯片的结构图, DMFB生物芯片由两个分离的板组成, 底板由一个二维(2D)电极阵列组成, 顶板由接地电极组成, 底板和顶板表面由一层绝缘固体和一层疏水层组成。

图 1(b)所示是 DMFB 生物芯片执行生化协议的基础原理, 即介质上电湿润原理(electrowetting-on-dielectric, EWOD)。DMFB 电极操纵离散液滴在两板之间分裂、混合, 液滴最初静止在疏水表面上, 对液滴下底板和绝缘层之间的电极施加电势, 液滴与底板的夹角 θ 减小, 造成了静止液滴向两侧分裂的现象。EWOD 原理可用 Lippmann-Young[18]方程建立模型:

$$\cos\theta(V)=\cos\theta(0)+\left(\epsilon_0\epsilon_r/2dY_{LG}\right)V^2 \qquad (1)$$

其中 V 是两块平板间的电压, θ 是未施加电压时的平衡接触角, ϵ_0 是真空中的介电常数, ϵ_r 是底部绝缘体的介电常数, d 是其厚度, Y_{LG} 是气体和液体界面的张力。



(a) 结构侧视图
(a) Side view of structure

(b) 介质上电湿润原理
(b) Principle of EWOD

图 1　结构侧视图与介质上电湿润原理示意图

Figure 1　Side view of the structure and schematic of electrowetting on dielectrics

2.2　DMFB 网络物理系统

图 2 以实时定量聚合酶链反应(Quantitative real-time polymerase chain reaction, RT-qPCR)为例, 说明 DMFB 在网络物理系统的支持下完成自动化设计、执行、控制和分析生化协议的全过程, 主要包含第三方制造商和外包设计生产 DMFB 以及终端用户使用 DMFB 平台进行生化协议测定, 其中 DMFB 及网络物理系统构建用户数据平台, 包括生物芯片、传感器和控制中心。假设数据库的属性可分类为隐私属性和公共属性, 隐私属性需要对其进行处理后发布, 公共属性则可以直接公布。然而, 隐私属性和公共属性不存在明显分界, 并且任何属性组合都有可能泄露个人信息。这个结论尤其符合 DMFB 网络物理系统环境。

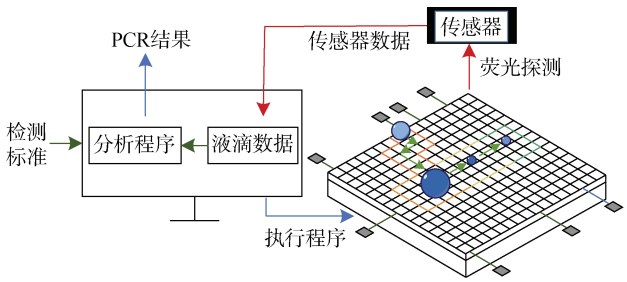


图 2　基于 DMFB 的实时定量聚合酶链反应

Figure 2　RT-qPCR based on a DMFB

2.3　增强隐私保护的用户数据安全平台

图 3 所示为增强隐私保护的生化协议执行平台, 其中 DMFB 生物芯片执行生化协议并管理用户数据, 划分为数据收集和数据发布两个阶段。液滴测定结果、传感器的校准范围作为生化协议的实时数据传输到 DMFB 控制中心, 控制器启动 DMFB 错误检测机制并返回反馈信号, 决定生化协议执行方向。测定结束后控制器存储过程数据及最终结果。在增强隐私保护的执行方案中, 若干次测定结果组成聚合信息, 等待上传到云端数据库中。图 3 使用聚合而来的用户数据, 经由隐私模型处理, 发布的数据具有可计量的隐私保障, 研究人员可以根据各自的研究目的去使用公共数据集。

2.4　差分隐私技术

本节介绍差分隐私的定义、实现机制和组合性原理。

2.4.1　定义及性质

表 3 中总结了差分隐私常用符合及解释。

定义 1: 对数据集 D 的各种映射函数 f 为查询。

定义 2: 相邻数据集为数据集 D 和数据集 D' 相差一条信息 $x_i \in N^{|X|}$, 即汉明距离 $d(D, D')=1$.

定义 3: 对任意相邻数据集 D 和数据集 D' 及任意算法结果 S, 差分隐私机制 K 如果满足:

$$P[K(D)\in S]\leqslant \exp(\varepsilon)\times P[K(D')\in S]+\delta \tag{2}$$

则差分隐私机制 K 是 ε 差分隐私[22], 其中, ε 是隐私预算, exp 是指数函数, δ 是失败概率, 概率 $P(K)$是抛硬币 K 次的概率。当 $\delta=0$, ε 差分隐私具有严格的定义; 当 $\delta>0$, (ε, δ)差分隐私提供了违反严格 ε 差分隐私的自由。

定义 4: 数据集 D 的查询 f: $D\rightarrow \boldsymbol{R}^k$, $\boldsymbol{R}^k$ 为 k 维实数向量, 是 f 返回的查询结果, 对于任意一对相邻数据集 D 和 D', f 的全局敏感度为:

$$GS_f=\max_{D,D'}\left\|f(D)-f(D')\right\|_1 \tag{3}$$

对于给定数据集 D 和 D', f 的局部敏感度为:

$$LS_f=\max_{D,D'}\left\|f(D)-f(D')\right\|_1 \tag{4}$$

其中$\|\cdot\|_1$ 是一阶范数, 全局敏感度与数据集无关, 只与查询结果有关。

2.4.2　实现机制

ε 差分隐私可以通过拉普拉斯机制和随机响应机制实现, (ε, δ)差分隐私可以通过高斯机制实现。

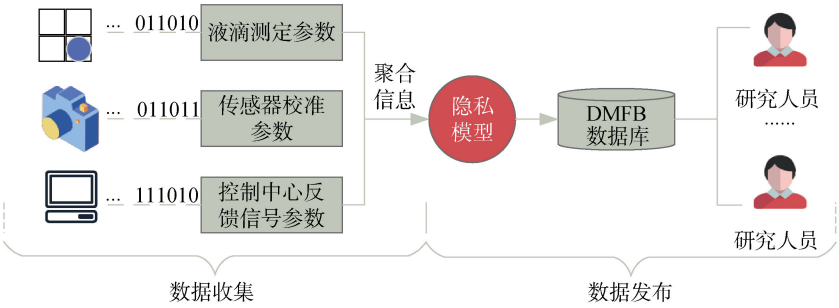


图 3 增强隐私保护的用户数据安全平台

Figure 3 Enhanced privacy protection DMFB's user data security platform

表 3 差分隐私常用符号和解释

Table 3 Notations and meanings for differential privacy

符号	解释	符号	解释
D	数据集	δ	失败概率
GS_f	f的全局敏感度	LS_f	f的局部敏感度
d	汉明距离	x_i	一条数据
$\boldsymbol{R}^k$	k 维度实数向量	K	差分隐私机制
f	查询	X	数据库空间
σ	方差	p	真实比例
ω	随机扰动	ε	隐私预算

1) 拉普拉斯机制

定理 1: 给定数值查询函数 f, 数据集 D 和拉普拉斯扰动 $Y\sim\mathrm{Lap}\left(\frac{\Delta f}{\varepsilon}\right)$, 差分隐私机制 K 满足 ε 差分隐私[22]。

$$K_L(D,f,\varepsilon)=f(D)+\mathrm{Lap}\left(\frac{\Delta f}{\varepsilon}\right) \tag{5}$$

根据定理 1, 拉普拉斯机制 K_L 的扰动被缩放为 $\omega=\frac{\Delta f}{\varepsilon}$, 满足 ε 差分隐私定义。

2) 高斯机制

高斯机制的目标是降低达到同等隐私预算 ε 而添加的扰动量, 失败概率是 $1-\delta$, 其中 δ 是一个很小的数, 扰动量由问题敏感度 GS_f 和隐私预算 ε 决定。

定理 2: 给定数值查询函数 f, 隐私数据库 D 和高斯扰动 $N\sim N(0,\sigma)$, 如果 $\sigma^2=2s^2\log(1.25/\delta)/\varepsilon^2$, 则高斯机制 K_G 满足(ε,δ)差分隐私。

$$K_G(D,f,\varepsilon)=f(D)+N \tag{6}$$

高斯扰动可能与数据集中原本存在的扰动同分布[22], 因此更适合 DMFB 的传感器数据集。

3) 随机响应机制

回顾表 1, 用户查询关于 DMFB 的问题可以为“型号 D1100 的 DMFB 有多少条反馈信号(QA)为 1?”在随机响应机制中, 第三方不必获得真实答案, DMFB 会自行聚合扰动答案以逼近真实答案。DMFB 控制器对任意一个随机检查点的反馈信号视作一次独立随机事件, 结果为响应。抛掷一枚硬币, 正面向上则响应反馈信号真实值(1 或者 0), 反面向上则抛掷第二枚硬币, 第二枚硬币正面向上响应 1, 反面向上响应 0。反馈信号的随机响应都会以一定概率诚实回答真实信号值, 假设在数据库中“$QA=1$”的真实比例为 p, 那么扰动答案中“$QA=1$”的比例则是 $p'=1/2\,p+1/4$。对于以上扰动答案, 推测出“$QA=1$”的比例是 $2\,p-1/2$, 因此查询结果为 $2\,p-1/2$。

令单条查询函数为 f, 对于任意 $x_i\in D$ 和随机响应机制为 $K_R(D,f,\varepsilon)$, 满足下式[22]:

$$\ln\frac{P(x=0\,|\,r=0)}{P(x=1\,|\,r=0)}=\frac{3/4}{1/4}=\ln 3 \tag{7}$$

同理,

$$\ln\frac{P(x=1\,|\,r=1)}{P(x=0\,|\,r=1)}=\frac{3/4}{1/4}=\ln 3 \tag{8}$$

其中 r 为响应, 在二元分布中, 给定随机响应机制 K_R 如下:

$$K_R(D,f,\varepsilon)=\begin{cases} f(X),p \\ 1-f(X),1-p \end{cases} \tag{9}$$

则随机响应机制 K_R 满足 ε 差分隐私, 且概率 $p=\frac{e^{\varepsilon}}{1+e^{\varepsilon}}$。

2.4.3 组合性原理

1) 串行组合原理

给定数据集 D 以及一组关于 D 差分隐私算法 $K_1(D)$, $K_2(D),\cdots,K_m(D)$, 算法 $K_i(D)$分别满足 ε_i 差分隐私且任意两个算法的随机过程互相独立。则这些算法组合起来的算法满足 $\sum_{i=1}^{m}\varepsilon_i$ 差分隐私。

2) 并行组合原理

定义差分隐私算法所保护数据库集合 D 中的元

素 x 定义在集合 O 上, 令$\{O_1, O_2,\ldots, O_t\}$为 O 的一种划分, 将数据集D划分成不同的子集, 将满足划分子类 O_i 的数据子集为 D_i, 则 $D_i=D\cap R_i$。记 $K_1(D)$, $K_2(D)$,…, $K_m(D)$分别表示输入数据集为 D_1, D_2,…, D_t 的一系列满足 ε 差分隐私算法且任意两个算法的随机过程相互独立。则这些算法组合起来的算法也满足 ε 差分隐私。

3 基于 DMFB 生化协议参数的差分隐私方案

DMFB 执行生化协议过程中输出测定结果涉及到的变量为生化协议参数, 它包括测定结果参数、传感器校准范围参数和控制器反馈信号参数。

3.1 生化协议参数

液滴在 DMFB 平台上, 以样本和试剂为输入, 通过液滴分裂、混合和移动完成特定目标的样品制备阶段, 称为片段协议, 用符号 I 表示。 图 4 由 5 个片段协议 I 完成 RT-qPCR, 由 DMFB 控制电极输出测定结果 p, 连接 4 个传感器, 其读数为 s, DMFB 控制中心对输出的测定结果 p 判断后返回反馈信号 QA。

3.2 测定结果隐私化

研究人员在 DMFB 生物芯片上多次实验得到片段协议 I_i 的测定结果参数 p_i 有效范围为 $p_i\in[v_{i\min}, v_{i\max}]$。参数值 p_i 只有在满足有效范围$[v_{i\min}, v_{i\max}]$的情况下才会被控制电极输出, 作为样品制备算法的结果[28]。表 4 总结了测定结果参数隐私化算法使用的符号及解释。

3.2.1 隐私化原理

设第 i 测定结果参数 p_i, 其取值范围 $p_i\in[v_{i\min}, v_{i\max}]$。假设测定结果参数 p_i 精确度为 c_i, 则 p_i 可取的有效离散值的总数为 $N_{i\text{val}}$。

$$N_{i\text{val}}=\left(v_{i\max}-v_{i\min}\right)/c_i \tag{10}$$

有效离散值总数 $N_{i\text{val}}$ 为参数 p_i 可取值的总数, $N_{i\text{val}}$ 的二进制长度 l 不超过$\left\lceil \log_2\left(N_{i\text{val}}\right)\right\rceil$, 取整函数用于确保每个可能的参数值映射到至少一个二进制表示形式。设随机扰动 ω_i 是一个随机的 l 位二进制。

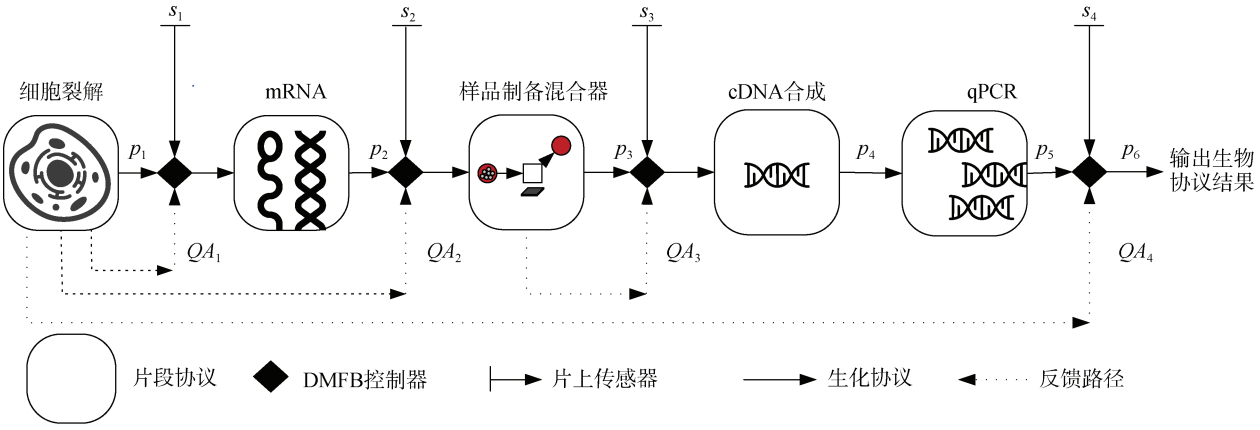


图 4 DMFB 片段协议及生化参数示意图

Figure 4 Fragment protocol and parameter diagram of the DMFB

表 4 符号及解释

Table 4 Notations and meanings

符号	解释
p_i	第 i 参数值
$[v_{i\min}, v_{i\max}]$	p_i 有效范围
c_i	p_i 精确度
$N_{i\text{val}}$	p_i 可取的有效离散实值的总数
ω_i	高斯机制扰动集合
l	ω_i 的最大二进制长度
s	拉普拉斯机制敏感度
ε	拉普拉斯机制隐私预算
p'_i	最接近 p_i 且 c_i 正确的 p_i 修正值

$$\omega_i = N_{i\text{val}} + \text{Lap}\left(\frac{s}{\varepsilon}\right) \tag{11}$$

其中 $l\leqslant\left\lceil \log_2\left(N_{i\text{val}}\right)\right\rceil$。根据返回的 $N_{i\text{val}}$, ω_i 定义满足 ε 差分隐私, 公式(12)将随机扰动 ω_i 嵌入到参数测定结果 p_i 中。

$$p_i=v_{i\min}+2^{-l}\text{int}\left(\omega_i\right)N_{i\text{val}} \tag{12}$$

函数 int(ω_i)将二进制随机 ω_i 转换为无符号整数表示。由公式(12)计算的参数测定结果 p_i 可能超出精确度 c_i, 为此, 将 p_i 修正为最接近且精确度正确的值 p_i'。

$$p'_i=\left\lceil p_i/c_i\right\rceil\times c_i \tag{13}$$

3.2.2 隐私化算法

算法 1.测定结果参数隐私化算法。

输入: 真实 p_i

输出: p'_i

过程 1: 计算 $N_{i\text{val}}$

过程 2: 生成随机扰动 ω_i

令 Lap(s/ε)中敏感度 $s = 1$

FOR ε IN range (0, 1):

计算 ω_i

IF $l = \text{len}(\omega_i) \leqslant \text{ceiling}(\log_2(N_{i\text{val}}))$ THEN

$\omega_i = \text{int}(\omega_i)$

计算 p_i

END IF

过程 3: 修正 p_i, 计算 p'_i

文献[13]中的免疫分析方案报告了在 1hz 时钟下运行的生物芯片上的孵化时间 t_1 在[360, 600]范围内, $c_1 = 1$, 有效孵化时间总数 $N_{1\text{val}} = (600-360) / 1 = 240$, 即 t_1 的有效离散值总数为 240 个。设 ε 差分隐私预算 $\varepsilon = 0.1$, 对于参数 t_1, 嵌入的拉普拉斯扰动长度不超过 $l \leqslant \lceil \log_2(N_{i\text{val}}) \rceil = \lceil \log_2(240) \rceil = 8$ 位。设 ω_1= '11110111', 则含有差分扰动培养时间 t_1 为 591.90 s, 修正至精确度为 $t'_1 = 592$ s。

3.3 校准范围隐私化

DMFB 控制器实时配置 DMFB 以进行差错恢复, 其中传感器读数 s_i、精确度 c_i 和校准范围[$v_{i\min}$, $v_{i\max}$]存储在 DMFB 数据库。高斯机制的生成随机扰动[$\omega_{i\min}$, $\omega_{i\max}$], 将 $\omega_{i\min}$ 和 $\omega_{i\max}$ 分别嵌入 $v_{i\min}$ 和 $v_{i\max}$ 得到[$v'_{i\min}$, $v'_{i\max}$]。表 5 总结了校准范围参数隐私化算法使用的符号及解释。

表 5 符号及解释

Table 5 Notations and meanings

符号	解释
s_i	传感器 i 的读数 s_i
$s_{i\text{tol}}$	第 i 传感器相同精确数的容忍值
n	数据集的大小
c_i	传感器数据精确度
ε	高斯机制隐私预算
δ	高斯机制失败概率
s	高斯机制敏感度
σ	高斯(正态)分布抽样方差
[$v_{i\min}$, $v_{i\max}$]	传感器 i 的校准范围
[$v'_{i\min}$, $v'_{i\max}$]	嵌入扰动的校准范围
ω_i	高斯扰动集合
l	ω_i 的最大长度

3.3.1 隐私化原理

定义 5: 传感器数据精确度为 c_i, 取 $c_i = c_i / 10$, 传感器 i 的读数 s_i 能够“容忍的”每一个真实值叫做 s_i 的容忍值, $s_{i\text{tol}}$。

$$\left|s_{itol} - s_i\right| \leqslant c_i/2 \tag{14}$$

设传感器 i 的读数 s_i 的容忍值 $s_{i\text{tol}}$ 组成 s_i 数据集, 数据集的大小为 n, 差分隐私失败概率为 $\delta = 1/n^2$。如下定义 ω_i 满足了(ε, δ)差分隐私:

$$\omega_i = s_{i\text{tol}} + N\left(\sigma^2\right) \tag{15}$$

其中 $\sigma^2 = \dfrac{2s^2 \log \dfrac{1.25}{\delta}}{\varepsilon^2}$, 而 $N(\sigma^2)$表示从中心为 0 且方差 σ 的高斯(正态)分布抽样, s、δ、ε 分别是高斯机制敏感度、失败概率和隐私预算。公式(15)生成了高斯扰动集合 ω_i, 因此存在 $\omega_{i\min}, \omega_{i\max} \in \omega_i$, 使得$|\omega_{i\min} - v_{i\min}|_{\min}$, $|\omega_{i\max} - v_{i\max}|_{\min}$ 分别成立, 其中$|\omega_i| \leqslant \left\lfloor \left(\left\lceil \log_2\left(N_{i\text{val}}\right) \right\rceil - 1 \right) / 2 \right\rfloor$。

设传感器 i 的读数 $s_i \in [v_{i\min}, v_{i\max}]$, 其容忍值 $s_{i\text{tol}}$ 的范围是 $s_{i\text{tol}} \in [-c_i / 2 + v_{i\min}, c_i/2 + v_{i\max}]$, 换言之, 通过分别在 $v_{i\min}$ 和 $v_{i\max}$ 添加扰动 $\omega_{i\min}$ 和 $\omega_{i\max}$ 修改传感器读数 s_i 的校准范围[$v_{i\min}$, $v_{i\max}$], 如下所示:

$$\begin{aligned} v'_{i\min} &= v_{i\min} - \frac{c_i}{2} \cdot \text{int}\left(\omega_{i\min}\right) \\ v'_{i\max} &= v_{i\max} - \frac{c_i}{2} \cdot \text{int}\left(\omega_{i\max}\right) \end{aligned} \tag{16}$$

int(ω)将二进制数 ω 转换为无符号整数形式, 校准范围参数隐私化算法将校准范围参数修改为[$v'_{i\min}$, $v'_{i\max}$]。

3.3.2 隐私化算法

算法 2. 校准范围参数隐私化算法.

输入: $v_{i\min}$, $v_{i\max}$, c_i

输出: $v'_{i\min}$, $v'_{i\max}$

过程 1: 对 $v_{i\min}$, $v_{i\max}$ 求出所有容忍值

FOR *tmp* IN [$v_{i\min}$, $v_{i\max}$]:

WHILE(1)

{

IF round($s_{i\text{tol}} - c_i \times 0.1$) = *tmp* THEN

$s_{i\text{tol}} = s_{i\text{tol}} - c_i \times 0.1$

ELSE IF round ($s_{i\text{tol}} + c_i \times 0.1$)= *tmp* THEN

$s_{i\text{tol}} = s_{i\text{tol}} + c_i \times 0.1$

IF END

}

END WHILE

ED FOR

过程 2: 高斯扰动参数计算

计算 s_i 的有效离散值总数 $N_{i\text{val}}$

计算数据集大小 $n = \text{len}(s_{i\text{tol}}) \times N_{i\text{val}}$

计算失败概率 $\delta = 1/n^2$

令高斯机制的敏感度 s=1, 计算方差 σ, 计算公式为 $\sigma=\sqrt{\dfrac{2s^2\log\dfrac{1.25}{\delta}}{\varepsilon^2}}$。

过程 3: 输入数据集 $s_{i\text{tol}}$ 生成随机扰动 ω_i

过程 4: $\omega_{i\min} = \min(\omega_i)$, $\omega_{i\max} = \max(\omega_i)$

过程 5: 计算传感器校准范围$[v'_{i\min}, v'_{i\max}]$

读数为 s_1 精确度 c_1 = 0.001, s_1 校准范围为$[v_{1\min}, v_{1\max}]$=[0,0.02]。当 $v_{1\min}$ = 0 时, 容忍范围 $s_{1\text{tol}} \in [-0.004, 0.004]$, 当 $v_{1\max}$ = 0.02, 容忍范围 $s_{1\text{tol}} \in [0.015, 0.024]$, 解释为在相同情形下传感器读数为 0.02 的容忍范围是[0.015, 0.024]。因此当校准范围在 [0, 0.02], s_1 的容忍范围 $s_{1\text{tol}} \in [-0.004, 0.024]$, 根据 s_1 数据集的定义得到 n, 故高斯机制的失败概率 $\delta = 1/n$, 隐私预算 ε 取值(0.0227,1)。根据公式(15)产生一个集合 ω_1, 根据公式(10)可知 s_1 的有效离散值总数 $N_{1\text{val}} = (v_{1\max} - v_{1\min}) / c_1$=0.02/0.001=20, 因此集合中元素最大长度不超过 $l \leqslant \left\lfloor \left(\left\lceil \log_2(20) \right\rceil - 1 \right) / 2 \right\rfloor$。满足 $|\omega_{1\min} - 0|_{\min}$, $|\omega_{1\max} - 0.02|_{\min}$ 的 $\omega_{1\min}$ 和 $\omega_{1\max}$, 分别是 0.207, 7.609, 根据公式(16)将高斯扰动 $\omega_{1\min}$ 和 $\omega_{1\max}$ 分别嵌入[0,0.02]后校准范围为[0,0.023]。

3.4 反馈信号隐私化

图 6 中 QA 即控制器对 DMFB 发出的反馈信号, 由随机响应算法, 得到DMFB 的随机响应结果 R, 对 R 进行编码扰动后为 B。表 6 总结了校准范围参数差分隐私计算中使用的符号及解释。

表 6 符号及解释

Table 6 Notations and meanings

符号	解释
q_i	随机检查点
$E_{\text{threshold}}$	q_i 错误阈值
e	累计误差率
n	随机检测点数量
QA_i	在 q_i 的反馈信号
R_i	QA_i 的随机响应结果
FR	$QA_i = 0$ 且 $R_i = 1$
TT	$QA_i = 1$ 且 $R_i = 1$
T	$R_i = 1$
B	R 的扰动结果
ε	随机响应机制的隐私预算
p, q	隐私预算的参数

3.4.1 隐私化原理

图 5 为反馈信号示意图, 其中液滴测定参数 $p_i \in [v_{i\min}, v_{i\max}]$, 传感器 i 读数 $s_i \in [v_{i\min}, v_{i\max}] \subseteq [v'_{i\min}, v'_{i\max}]$, DMFB 的随机检查点 $q_i \subseteq p_i \in [v_{i\min}, v_{i\max}]$, 其中$[v_{i\min}, v_{i\max}]$存储在控制器中, $[v'_{i\min}, v'_{i\max}]$由 DMFB 实时生成。定义 DMFB 控制器对 q_i 的反馈信号为质量评估参数 QA_i, DMFB 控制器根据 q_i 是否满足原范围$[v_{i\min}, v_{i\max}]$返回 QA_i, DMFB 接收 QA_i 并决定是否重新执行片段协议以启动差错恢复机制。

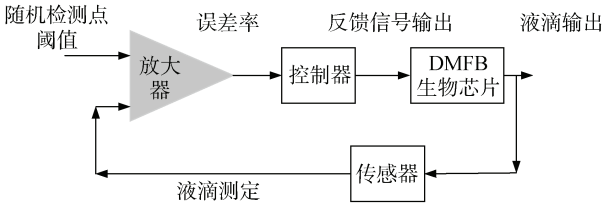


图 5 反馈信号示意图[24]

Figure 5 Schematic diagram of feedback signals[24]

$$QA_i = \begin{cases} 1, v_{i\min} \leqslant q_i \leqslant v_{i\max} \\ 0, \text{其他情况} \end{cases} \tag{17}$$

其中 i=1,2,...,n, 定义随机响应机制对于 QA_i 的响应为 R_i, R_i 的取值范围{0,1}, 由于抛掷一枚硬币正面向上的概率是 1/2, $QA_i = 1$ 且 $R_i = 1$ 的概率是 1/2, R_i 有 1/4 的概率随机取 1 和 1/4 的概率随机取 0, 由此当响应 $R_i = 1$ 时, $QA_i = 0$ 成立的概率为 1/4。定义符号 FR, 代表响应 $R_i = 1$ 且反馈信号 $QA_i = 0$, 易得:

$$FR = \text{sum}(R_i) / 4 = n / 4 \tag{18}$$

其中 sum(R_i)函数是计算 R_i 的总数。设 $R_i = 1$ 的数量为 T, 计算公式为:

$$T = \text{sum}(R_i = 1) \tag{19}$$

其中 i=1,2,...,n。定义符号 TT, 代表响应 $R_i = 1$ 且反馈信号 $QA_i = 1$, 由 T 和 FR 可以得到 TT 计算如下:

$$TT = 2(T - FR) \tag{20}$$

3.4.2 隐私化算法

随机响应机制[25]的隐私预算 ε 由概率 p 和 q 共同决定。

$$\varepsilon = \log\left[p(1-q) / (1-p)q \right] \tag{21}$$

算法 3.随机响应算法.

输入: QA

输出: R

过程 1: 计算 QA 的响应集 R

FOR QA_i IN QA:

令 $truthful\_ R_i = QA_i$

IF random(0, 2) == 0//抛掷一枚硬币正面向上

$R_i = truthful\_ R_i$

ELSE//否则抛掷第二枚硬币

R_i = random(0, 2) == 0//随机回答 1 或 0

END IF

END DEF
END FOR
过程 2: FR = sum(R)/4 // R_i = 1 且 QA_i = 0
过程 3: T = sum(R_i = 1) // R_i = 1
过程 4: TT = sum(R_i = 1) – FR//R_i = 1 且 QA_i = 1
TT = 2TT //1/2 概率 R_i 随机

算法 4: 编码扰动算法[26].
输入: R, p, q
输出: R 的扰动结果 B
p = random(0, 1)
q = 1– p
FOR R_i IN R:
  IF R_i = 1
    IF random(0, 1) < p
      $B[i]$ = 1
      ELSE $B[i]$ = 0
  ELSE IF R_i = 0
    IF random(0, 1) ⩽ q
      $B[i]$ = 1
      ELSE $B[i]$ = 0
  END IF
END FOR

4 安全性分析

本章介绍攻击模型并建立参数安全发布算法以及隐私安全性和数据可用性的指标体系，最后总结差分隐私保护技术在 DMFB 的安全应用场景。

4.1 攻击模型

差分隐私保护技术在 DMFB 的安全应用场景有生物实验室以及远程访问，根据设备是否在线分类为离线环境和在线环境，图 6 为该场景分类下的攻击模型示意图，其中攻击者可能是实验室内部人员和通过网络发起攻击的远程服务或者恶意软件。通过伪装液滴攻击 DMFB 生化协议平台，从而伪造和篡改生化协议；通过网络接口对运行时的 DMFB 进行拍摄，以盗取 DMFB 运行时液滴状态；甚至通过网络操纵 DMFB 的控制软件，篡改液滴运行状态或者传感器读取的数据，导致 DMFB 生化协议测定有误。攻击者以隐秘且无法追踪的方式对 DMFB 进行攻击，液滴运行数据是平台上的用户数据，传感器和控制中心都是获取用户数据的辅助设备。这一过程中，生物芯片设计者是防御者，用户信任生物芯片平台是最终目标。

4.2 参数安全发布算法

生化协议参数的安全发布算法引入如下所述的三个隐私化策略:

(1) 创建层级敏感度 HS, 根据用户身份 u 选择对应的层级敏感度 HS_u。

(2) 创建路由集合 $R_i(i = 1, \ldots, \eta)$, 对路由 R_i 上的单元阵列 g 上的参数信息 $T(g)$添加随机干扰。

(3) 创建路由交叉点 g_c 参数信息集合 $T(g_c)$, 根据权重对参数信息 $T(g_c)$添加随机干扰。

表 7 总结了参数安全发布算法中使用的符号及解释。

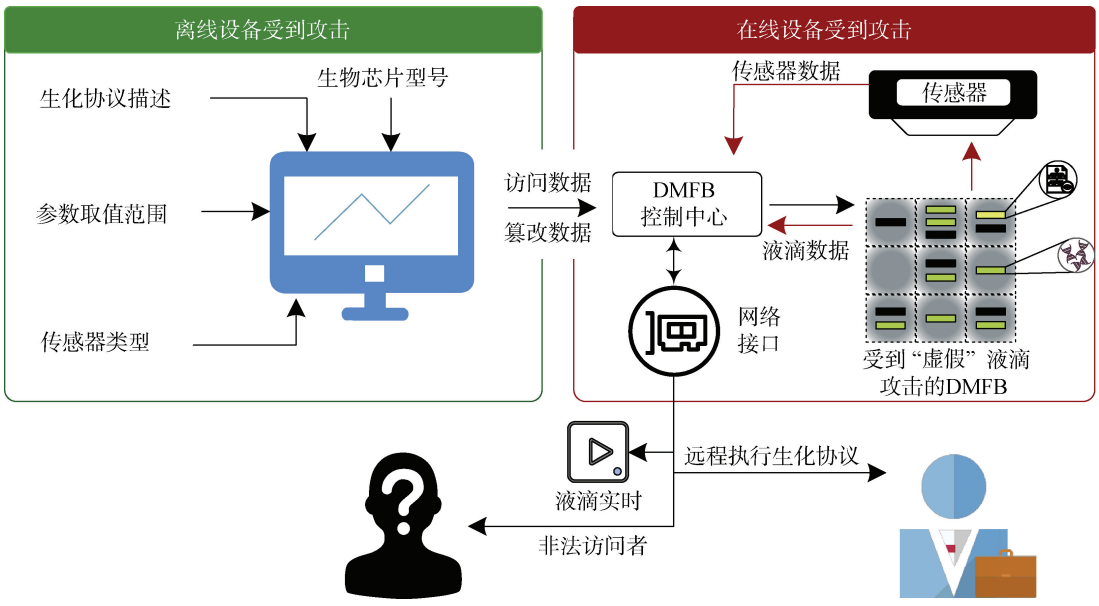


图 6 攻击模型示意图

Figure 6 Schematic diagram of attack model

表 7 符号及解释

Table 7 Notations and meanings

符号	解释
$g=(x,y)$	单元阵列坐标信息
R	g 的有序集合组成液滴路由
$D=\{R_1,\cdots,R_\eta\}$	数据集 D
D'	满足差分隐私的数据集 D
η	路由的数量
$T(g)$	参数信息集合
ε	隐私预算
g_c	路由交叉点
$G(g_c)$	路由交叉点集合
$W(g_c)$	路由交叉点权重集合
HS	层级敏感度
$d(g_i,g_j)$	单元阵列 g_i 和 g_j 之间的距离
u	用户身份
p	测定结果参数, 简称参数
s	传感器
QA	随机检查点反馈信号

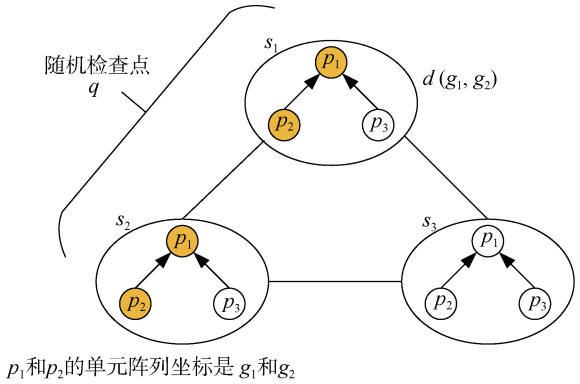


图 7 层级敏感度

Figure 7 Hierarchy sensitivity

4.2.1 层级敏感度

层级敏感度概念中只考虑 DMFB 数据库中这三类参数, 及 p_i, s_i 和 QA。定义: 查询的层级敏感度 L 为:

$$HS = HS_L = \max_L d\left(g_i, g_j\right) \tag{22}$$

其中, $d(g_i,g_j)$为单元阵列 g_i 和 g_j 在 DMFB 上的距离 L, 计算公式如下。

$$d\left(g_i, g_j\right) = \sqrt{\left(x_i - y_i\right)^2 + \left(x_j - y_j\right)^2} \tag{23}$$

如图 7 所示层次敏感度由 p, s 和 QA 组成。传感器 s_1, s_2 和 s_3 能够测量的液滴分别为 p_1, p_2 和 p_3, p_1, p_2 和 p_3 距离 L 分别为 $d(g_1,g_2)$, $d(g_2,g_3)$和 $d(g_1,g_3)$。当敏感度的层级为 p, 意味着液滴测定结果 p_i 会加入干扰, 又因为固定 p_i, 有且仅有一个 g_i 与之对应, 因此此时敏感度为 $HS_{pi}=d(g_i,0)$; 当敏感度的层级为传感器 s, 意味着传感器校准范围参数会加入干扰, 如图 7 所示, 当敏感度层级为 s_1, 传感器 s_1, s_2 和 s_3 能够测量的液滴分别为 p_1, p_2 和 p_3, 对 s_1 校准范围$[v_{1\min},v_{1\max}]$ 添加干扰, 可得 p_1, p_2 和 p_3 带有干扰, 此时敏感度为 $HS_s=\max\{d(p_1,p_2), d(p_2,p_3), d(p_1,p_3)\}$; 如图 7, 随机检查点经过的路径可组成液滴路由 R_q, 由于可能经由多个传感器 s, 因此此时敏感度 $HS_{QA}=\{d(s_i,s_j)\}$, 图 7 中敏感度 $HS_{QA}=d(s_1,s_2)$。

4.2.2 路由交叉点

定义 1: 以 DMFB 生物芯片左下角为原点, 以水平方向为 x 方向, 以垂直方向为 y 方向, 建立如图 8 所示直角坐标系, 则生物芯片的单元阵列为 $g=(x,y)$。

定义 2: 一组有序的单元阵列坐标集合为路由 R, g_{in} 和 g_{ou} 分别是路由 R 的首尾坐标。图 8 中路由 R_1={(1,2),(2,2),(3,2),(4,2),(5,2),(5,3),(5,4),(5,5),(5,6),(5,7),(5,8),(6,8), (7,8), (8,8)}, 其中 g_{in}=(1,2), g_{ou}=(8,8); 路由 R_2={(1,6),(2,6),(3,6),(4,6),(5,6),(6,6),(7,6),(7,5), (7,4),(7,3),(7,2),(8,2)}, 其中 g_{in}=(1,6), g_{ou}=(8,2)。

定义 3: $R_i \cap R_j=g_c$, g_c 称为 R_i 和 R_j 的路由交叉点, i, $j=1,\ldots,\eta$ 且 $i\neq j$, η 为路由的数量。

定义 4: 路由交叉点 g_c 对于路由 R_i 的权重 $w(g_c, R_i)$计算如下:

$$w\left(g_c, R_i\right) = \frac{d\left(g_c, g_{in}\right)}{d\left(g_{in}, g_{ou}\right)} \tag{24}$$

其中, j=1, …, n, R_i 是数据集 D 第 i 条液滴路由, g_c 是路由 R_i 的交叉点, g_{in} 和 g_{ou} 是路由 R_i 的入口和出口的坐标信息。图 8 中 g_c= (5,6)是路由 R_1 和 R_2 的路由交叉点, (5,6)的参数信息集合 $T(g_c)=\{p_1, p_2\}$解释为单元阵列(5,6)上液滴两次测定结果参数为 p_1, p_2, 根据定义, $w((5,6), R_1)$=0.613, $w((5,6), R_2)$=0.433。

4.2.3 算法描述

算法 5. 参数安全发布算法.

输入: D, ε, u

输出: D'

过程 1: 将隐私预算分为 $\varepsilon/2$ 和 $\varepsilon/2$

过程 2: 根据用户 u 选择层级敏感度 HS_u[策略 1]

过程 3: 创建交叉点坐标集合 $G(g_c)$及其权重集合 $W(g_c)$

FOR $T(R_i)$IN $T(D)$://路由参数信息隐私预算是 $\varepsilon/2$

过程 4: $T(R_i)$分配 $\varepsilon/2$ 隐私预算 [策略 2]

END FOR

FOR $T(g_c)$ IN $T(D)$:

过程 5: $T(g_c)$ 按照权重分配隐私预算 $\varepsilon_i=(1-w_i/\sum w_i)\ \varepsilon/2$[策略 3]

END FOR

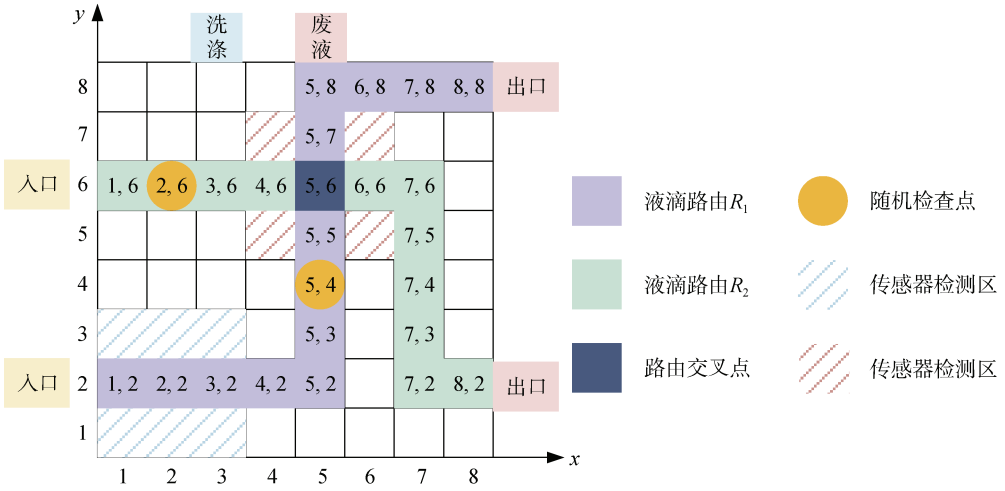


图 8　液滴路由及路由交叉点

Figure 8　Droplet routing and routing intersection

数据库 D 由 R_1, …, R_η 组成, 且用<g, $T(g)$>的有序对的形式存储, 其中 $T(g)$是 g 点上所有测定结果参数的参数信息集合。路由 R_i 的查询作用于数据集 D, 分配了 $\varepsilon/2$ 隐私预算, 路由交叉点 g_c 的查询作用于路由 R_i, 根据 R_i 上 g_c 权重分配隐私预算。过程 4 参数安全发布算法中策略 2 和策略 3 的隐私预算分配见表 8, 算法 5 满足 ε 差分隐私性质。

表 8　参数安全发布算法中隐私预算分配

Table 8　Privacy budget allocation in algorithm5

	隐私预算
路由参数信息	$\varepsilon/2$
交叉点参数信息	$(1-w_i/\sum w_i)\ \varepsilon/2$

4.3　篡改攻击

攻击者试图篡改液滴浓度、孵化时间、混合时间等测定结果参数以及传感器校准范围导致测定误差。定义测定结果和传感器校准范围被篡改成功的概率为篡改概率 PT, 则抗篡改概率为 $1-PT$。

4.3.1　测定结果参数

测定结果参数 p_i 的篡改值满足$[v_{i\min}, v_{i\max}]$时攻击成功。当 $p_i \pm \Delta p$ 恰好取值 $v_{i\min}$ 或 $v_{i\max}$, 此时称 Δp 为 p_i 可篡改范围。图 9(a)所示是 p_i 一种可行的扰动嵌入方案。$|p_i - v_{i\min}| \leqslant |v_{i\max} - p_i|$时, p_i 和 p'_i 可篡改范围 Δp 分别为$|p_i - v_{i\min}|$和$|p'_i - v_{i\min}|$, 差分隐私保护缩小了$|p_i - p'_i|$的可篡改范围。当$|p_i - v_{i\min}| \geqslant |v_{i\max} - p_i|$时同理可证。综上所述, 测定结果参数 $p_i \in [v_{i\min}, v_{i\max}]$嵌入差分扰动 $p'_i \in [v_{i\min}, v_{i\max}]$, 可篡改范围 Δp 缩小了$|p_i - p'_i|$, 计算篡改概率时采用$|v_{i\max} - p'_i|$和$|v_{i\min} - p'_i|$两者中较小值, 公式如下所示:

$$PT = \frac{\min(|v_i - p_i|)}{N_{i\text{val}}} \tag{25}$$

其中 $v_i \in \{v_{i\min}, v_{i\max}\}$, $i = 1,2, \ldots, k$, k 是 p_i 个数, $N_{i\text{val}}$ 是 p_i 的离散值总数, 取 k 个参数 PT 平均值作为测定结果参数的平均篡改概率 PT。

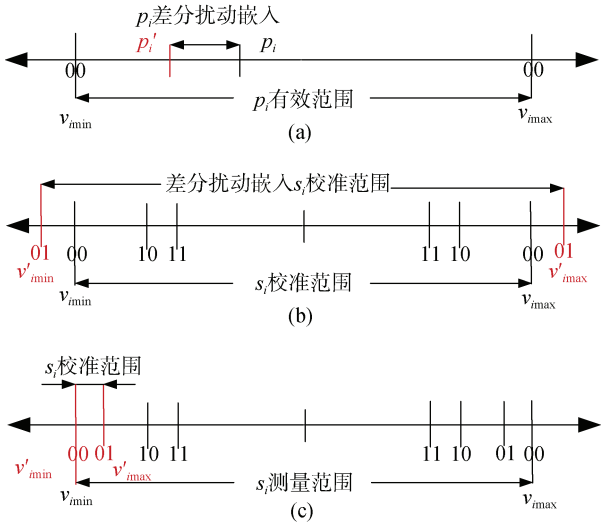


图 9　生化协议参数扰动嵌入

Figure 9　Biochemical protocol parameters and disturbance embedding

4.3.2　校准范围参数

设校准范围为$[v_{i\min}, v_{i\max}]$, 攻击者篡改校准范围$[v_{i\min}, v_{i\max}]$, 使 $v_{i\min}$ 减去不为 0 的数或 $v_{i\max}$ 加上不为 0 的数时篡改攻击成功。图 9(b)中生成高斯扰动 $\omega_{i\min}=|v'_{i\min} - v_{i\min}|$, $\omega_{i\max}=|v'_{i\max} - v_{i\max}|$并嵌入测量范围得到校准范围$[v'_{i\min}, v'_{i\max}]$, $[v'_{i\min}, v'_{i\max}]$为差分保护下的最大校准范围。图 9(c)所示, 存在 $v_i \in [v_{i\min}, v_{i\max}] \subseteq [v'_{i\min}, v'_{i\max}]$, $s_i = v_i \pm s_i$。攻击者以相同心理预期进行

对[$v_{i\min}$, $v_{i\max}$]篡改, 使 s_i 通过[$v'_{i\min}$, $v'_{i\max}$]而非[$v_{i\min}$, $v_{i\max}$]为篡改成功, 即攻击者的可篡改范围是 $v'_{i\max}-v_{i\max}+v'_{i\min}-v_{i\min}$。当 $v'_{i\min}\geqslant v_{i\min}$, $v_{i\max}\geqslant v'_{i\max}$, 此时可篡改范围是 $v_{i\max}-v'_{i\max}+v_{i\min}-v'_{i\min}$。综上所述, 校准范围的篡改概率 PT 为:

$$PT=\frac{\sum\left|v'_i-v_i\right|}{v_{i\max}-v_{i\min}} \tag{26}$$

其中 $v_i\in\{v_{i\min}, v_{i\max}\}$, $i=1,2,\ldots,m$, m 是 DMFB 片上传感器数量, 取 m 个传感器的篡改概率 PT 平均值作为传感器校准范围参数的平均篡改概率 PT。

4.4 数据可用性

对测定结果参数、校准范围参数和反馈信号参数的可用性度量指标进行研究, 分别定义置信分数 SC、校准度 E 和累计误差率 e。

4.4.1 测定结果参数

DMFB 生化协议测定参数存在合理范围, 即 $p_i\in[v_{i\min}, v_{i\max}]$。测定结果参数 p_i 是液滴浓度、孵化时间、混合时间等不同变量。对于测定结果 $p_i\in[v_{i\min}, v_{i\max}]$, 定义置信分数 SC 为攻击者对隐私保护之后的测定结果参数的数据可信度, 表示攻击者对测定结果参数的“信任”程度, 使用嵌入扰动前后的直线距离与离散值 $N_{i\text{val}}$ 的占比来衡量攻击者对参数 p'_i 的置信分数 SC。

$$SC=1-\frac{\left|p_i-p'_i\right|}{N_{i\text{val}}} \tag{27}$$

其中 $i=1,2,\ldots,k$, k 是 p_i 个数, $N_{i\text{val}}$ 是 p_i 的离散值总数, 取 k 个参数的置信分数 SC 平均值作为测定结果参数的平均置信分数 SC。

4.4.2 校准范围参数

在生化测定中获得尽可能高的校准度是必须的。测量血液样本中的葡萄糖浓度通过稀释剂浓度[21]和反应速率的校准曲线来度量此时葡萄糖变化量。DMFB 执行体外葡萄糖浓度监测, 避免了不正确的实验室操作导致的错误血液检测结果, 保证了用户的医疗数据安全。定义校准度 E 是经过差分扰动二进制嵌入后[$v'_{i\min}$, $v'_{i\max}$]对范围[$v_{i\min}$, $v_{i\max}$]的“覆盖”程度, 公式如下所示。

$$E=\frac{v'_{i\max}-v'_{i\min}}{v_{i\max}-v_{i\min}} \tag{28}$$

其中 $i=1,2,\ldots,m$, m 是 DMFB 片上传感器数量, 取 m 个传感器的校准度 E 的平均值作为传感器校准范围参数的平均校准度 E。

4.4.3 反馈信号参数

回顾图 6, 执行生化协议前, DMFB 控制器在生物芯片上部署随机检查点 q_i, 并存储 q_i 的错误阈值 $E_{\text{threshold}}$。反馈信号实际误差来自执行时反馈控制路径在控制器处的累计误差 e, 不应超过理论误差率 $E_{\text{threshold}}$。定义累计误差率 e 为随机检查点 q_i 的反馈信号 QA 误差。

$$e=\frac{\left|TT-T\right|}{T} \tag{29}$$

其中 TT 为响应 $R_i=1$ 且反馈信号 $QA_i=1$ 总数, T 为响应 $R_i=1$ 总数。

4.5 安全应用场景

对 DMFB 所面临的用户数据篡改攻击进行调查, 提出基于差分隐私的 DMFB 数据保护方案, 此外, 提出层次敏感度和分配隐私预应对查询输出的用户数据。如表 9 所示, 该方案可应用于生物实验室和在线(离线)DMFB 以抵御篡改攻击。使用差分隐私机制实现参数安全发布算法, 其应用环境可在实验室或者通过网络远程控制 DMFB, 由于安全发布算法不依赖于网络通信, 在离线环境下使用 DMFB 同样可以抵御篡改攻击, DMFB 数据发布存在数据隐私性和数据可用性之间的矛盾[27]。对其隐私安全性和数据可用性进行研究, 分别提出篡改概率 PT 和置信分数 SC、校准度 E、累计误差率 e, 在表 10 中对指标进行分类。

表 9 应用场景对比

Table 9 Comparison of application scenarios

	应用场景			篡改攻击
	实验室	在线设备	离线设备	
差分隐私保护	√	√	√	√
随机检测点[15]	√	√	√	√
逆向工程[9, 12-13]	√	√	×	√
数字版权管理[6-8]	√	√	√	×

表 10 指标分类

Table 10 Classification of indicators

	测定结果参数	校准范围参数	反馈信号参数
隐私安全性	篡改概率 PT		
数据可用性	置信分数 SC	校准度 E	累计误差率 e

5 实验及结果分析

模拟基于 DMFB 的免疫测定, 对 DMFB 片段协议测定结果参数和校准范围参数进行数据保护。

5.1 DMFB 免疫测定

DMFB 使用抗体包被的顺磁颗粒和外部磁场可以实现免疫测定的自适应。本节详细描述基于顺磁珠的 DMFB 免疫测定步骤。

图 11 使用了酶联免疫检测法, 包括病毒抗原的分离与纯化、病毒抗血清的制备和抗血清三个阶段。

第一步: 抗原会吸附在固相载体上, 因此为 DMFB 分配一个封装抗体附着的颗粒液滴, 并用磁铁将颗粒从稀释剂中分离出来(分离)。

第二步: 将可能含有特异性抗体的待测临床溶液与含抗原的颗粒混合 t_{Mx1} s 后孵育 t_{Inc1} s(混合)。

第三步: 用缓冲液洗涤步骤 2 孵育后的颗粒 t_{Wsh1} s(洗涤)。

第四步: 将酶标记抗体溶液与步骤 3 中颗粒混合 t_{Mx2} s 后孵育 t_{Inc2} s, 生成抗原-待测抗体-酶标记抗体的复合物(混合)。

第五步: 将步骤 4 中的复合物颗粒在洗涤缓冲液中洗涤 t_{Wsh2} s(洗涤)。

第六步: 将步骤 5 中的复合物颗粒从洗涤缓冲液中分离, 再与洗脱液混合 t_{Mx3} s, 并培养 t_{Inc3} s, 最后与显色增强剂溶液混合 t_{Mx4} s(混合)。

第七步: 将聚集的液滴孵育 t_{Inc4} s, 记录化学发光信号(检测)。

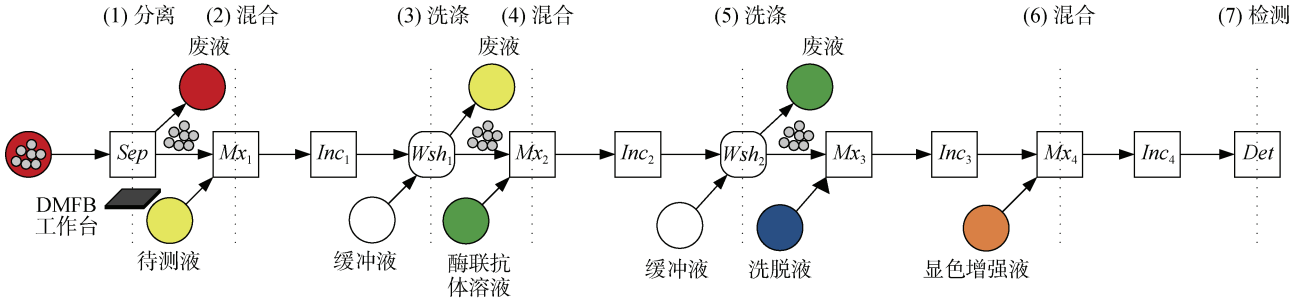


图 10 DMFB 实现免疫测定流程图

Figure 10 Flowchart of the DMFB implementation of immunoassay

5.2 DMFB 免疫测定仿真

仿真基于DMFB的免疫测定反应, 执行算法1～4, 对生化协议参数的隐私化算法的安全性和可用性进行评估。

5.2.1 测定结果参数仿真结果

表 11 是测定结果参数的仿真实验结果, 表中总结以十六进制表示的拉普拉斯扰动和扰动嵌入的结果及其置信分数 *SC* 和抗篡改概率 1−*PT*。置信分数 *SC* 和抗篡改概率 1−*PT* 分别平均达到 98.31%和 99.96%, 仿真结果表示, 基于差分隐私的 DMFB 数据保护方案为测定结果参数提供 98%的数据可用性和 99%的防篡改概率。

表 11 测定结果参数仿真实验结果

Table 11 Simulation experiment results of measurement result parameters

参数	扰动	参数值	修正值	置信分数(%)	抗篡改概率(%)
t_{Mx1}	0x165	431.3	432	99.80	99.61
t_{Inc1}	0x16b	435.6	436	99.70	100.00
t_{Wsh1}	0x3c	116.8	117	97.92	100.00
t_{Mx2}	0x38	173.1	174	96.50	100.00
t_{Inc2}	0x3c	176.9	177	96.40	100.00
t_{Wsh2}	0xd	72.9	73	96.25	100.00
t_{Mx3}	0x11	136.5	137	95.40	100.00
t_{Inc3}	0x38	172.8	173	95.03	100.00
t_{Mx4}	0xc7	196.1	197	98.06	100.00
t_{Inc4}	0x5a	73.9	74	95.60	100.00
平均				98.31	99.96

(注: 参数个数 $k=10$, 精确度 c_i=1.拉普拉斯机制的隐私预算 $\varepsilon=0.1$, 敏感度 s=1.)

5.2.2 校准范围参数仿真结果

表 12 是传感器校准范围参数的仿真实验结果, 原校准范围$[v_{i\min},v_{i\max}]$分别为[0,0.02]和[0,0.04], 高斯扰动范围分别在[000,111]和[0x0, 0x8], 此时校准度平均达到 100%, 抗篡改概率平均达 87.5%, 低于测定结果参数的抗篡改概率。

表 12 校准范围参数仿真实验结果

Table 12 Simulation experiment results of calibration range parameter

s_i	$[v_{\min},v_{\max}]$	$\omega_{i\min}$	$\omega_{i\max}$	$[v'_{i\min},v'_{i\max}]$	抗篡改概率(%)	校准度(%)
s_1	[0, 0.02]	000	111	[0, 0.023]	85	100
s_2	[0, 0.04]	0x0	0x8	[0, 0.044]	90	100
平均					87.5	100

(注: 传感器数 $m=2$, 精确度 $c_i=0.01$.高斯机制的隐私预算 $\varepsilon=0.1$,方差 $\delta=0.4\times10^{-4}$.)

5.2.3 反馈信号参数仿真结果

表 13 是控制器反馈信号参数的仿真实验结果, 在隐私预算[–4.39,4.39]范围内, QA_i=1 的预测数量(*TT*)平均为 209 个, 平均累计误差率为 7.56%。当隐私预算 ε=0.81 时, 最小累计误差率 e 为 2%, 当隐私预算 ε=0.4 时, 累计误差率 e 接近平均值, 为 8%。

表 13 反馈信号参数仿真实验结果

Table 13 Simulation experiment results of quality evaluation parameters

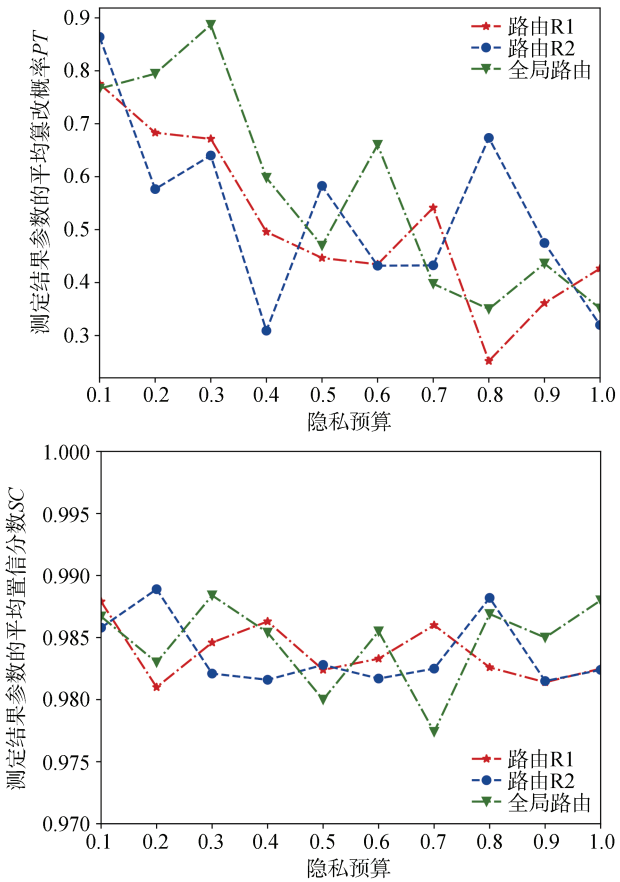
	隐私预算 ε	预测数量 *TT*	累计误差率 e(%)
	–4.39	264	32.00
	–2.77	212	6.00
	–1.69	224	12.00
	–0.81	166	17.00
	–0.40	244	9.00
	0.40	216	8.00
	0.81	204	2.00
	1.69	212	6.00
	2.77	210	5.00
	4.39	194	3.00
平均		209	7.56

(注: QA 的数据组是[(1, 200), (0, 1000)], 在随机响应机制中取隐私预算参数 $p=0.45$.)

5.3 算法性能验证

模拟 DMFB 环境, 执行算法 5, 在不同层级敏感度 *HS* 下对隐私预算 ε 进行研究, 调查此时算法的安全性和可用性。

图 11 为测定结果参数在算法 5 中的性能实验结果, 实验设置了路由 1、路由 2 和全局路由(路由 1 和路由 2)三组数据, 参数个数 k=10, 且层级敏感度 *HS* 为 p。对路由中的 $p_i(i=1,\ \ldots,\ k)$添加隐私预算 $\varepsilon\in[0.1,\ 1.0]$下生成的拉普拉斯扰动, 由算法 5 可知, 路由的隐私预算为 $\varepsilon/2$, 计算此时平均篡改概率 *PT*(如图 a)和平均置信分数 *SC*(如图 b)。在图(a)中, 除 ε=1.0 外, 全局路由的平均篡改概率 *PT* 高于路由 1 或路由 2; 三组路由平均篡改概率的下降区间分别为: [0.7,0.8], [0.5,0.7], [0.6,0.8], 最低平均篡改概率 *PT* 达 25%。在图(b)中, 隐私预算 $\varepsilon\in[0.1,1.0]$下三组路由的平均置信分数 *SC* 都达到了 98%, 全局路由的平均置信分数在隐私预算[0.5,1.0]内高于路由 1 和路由 2 的平均置信分数, 综合得隐私预算应取[0.4,0.8]。

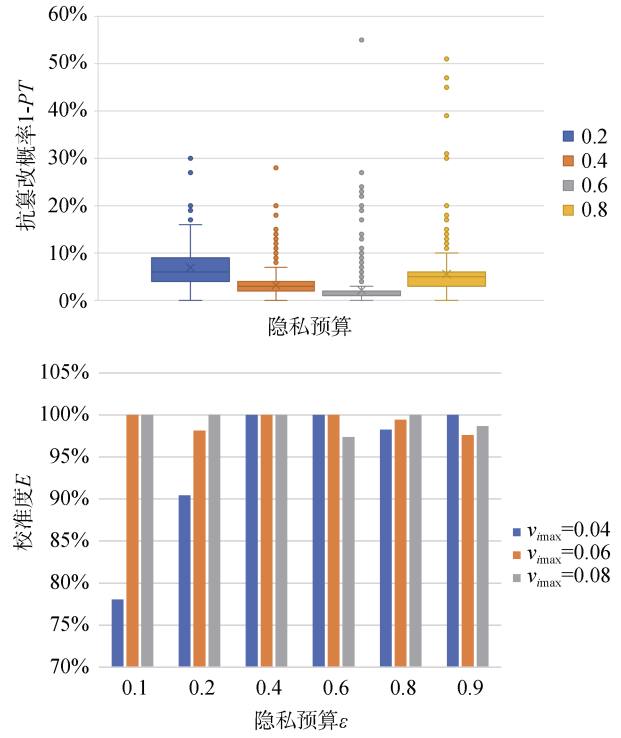


(注: 层级敏感度 *HS*=*p*, 参数个数 k=10, 拉普拉斯机制中隐私预算 $\varepsilon\in[0.1,\ 1.0]$)

图 11 测定结果参数(a)平均篡改概率(b)平均置信分数

Figure 11 Measurement result parameter (a) Average tampering probability (b) Average confidence score

图 12 为校准范围参数在算法 5 中的性能实验结果, 实验设置校准范围最小值 $v_{i\min}$= 0, 最大值 $v_{i\max}\in[0.2,\ 1.0]$, 传感器个数 m=5, 且层级敏感度 *HS* 为 s, 简单起见, 每一个传感器测量区域, 都设置相同数量和权重的路由交叉点。在图(a)中, 取 1-*PT* 作为观测, 隐私预算 $\varepsilon\in[0.2,\ 0.6]$ 时, 平均抗篡改概率 1–*PT* 高于[0.4, 0.8], 图中显示隐私预算最佳 $\varepsilon\in[0.4,$



(注: 层级敏感度 HS=s, 传感器个数 m=5, c_i=0.01 高斯机制中隐私预算 $\varepsilon \in [0.1, 1.0]$, 方差 δ= 0.4×10^{-4})

图 12 校准范围参数(a)平均抗篡改概率(b)平均校准度

Figure 12 Calibration range parameter (a) Average tamper probability (b) Average calibration degree

0.6], 此时抗篡改概率最高可以达到 99%。图(b)中, $\varepsilon \in [0.1, 1.0]$, 原校准范围[$v_{i\min}$, $v_{i\max}$]分别为[0, 0.04],[0, 0.06]和[0, 0.08]。结果表明平均校准度最低在 75%以上, $v_{i\max} \in [0.06, 0.08]$平均校准度最低 95%, 因此高斯扰动嵌入校准范围, 使最小值变小, 最大值变大, 一定程度上影响传感器校准度, 但是通过调整隐私预算可以提高校准度, 传感器校准范围最大/小值相差 0.04 或者更低时, 隐私预算 $\varepsilon \in [0.2, 0.9]$ 最低校准度达到 90%, 图中显示最佳隐私预算为 0.4。

5.4 算法有效性验证

本章对比文献[6]在测定结果和传感器校准范围上的隐私安全性以及模拟文献[17]中反馈控制路径, 与之对比数据可用性。

5.4.1 测定结果参数对比结果

表 14 所示为免疫测定协议的 10 个参数及篡改概率 PT 和置信分数 SC 分别使用文献[6]的 SHA3-256 哈希数字签名和拉普拉斯差分隐私机制对篡改攻击的抵御能力。在 10 个测定参数结果中, 相较于 SHA3-256 哈希数字签名, 拉普拉斯扰动的参数篡改概率平均下降 12.09%, 置信分数平均增加 0.1%。

5.4.2 校准范围参数对比结果

采用一组 DMFB 免疫测定中真实校准数据和模拟生成的其余三组测试用例, 对比传感器在文献[6]和高斯差分隐私机制下的抗篡改性能, 其中文献[6]用传感器最大(小)值加(减)二进制进行微调。用“DMFB Bioassay”作为数字签名生成 258 位的 SA3-256 哈希二进制串, 作为数字签名嵌入 DMFB 传感器校准范围, 为了实验对比明显, 数字签名避免选取‘000’。表 15 是高斯机制和数字签名分别对四组测试用例进行数据保护的结果。

表 16 是传感器篡改概率 PT 和校准度 E 对比结果, 在四个测试用例中, 本文方案传感器篡改概率平均减少 10.09%。其中, 在[0, 0.02]上篡改概率优化达到 15%, 在[0, 0.06]上数字签名为[11, 11], 高斯扰动为[–0x3, 0x6], 因此[0, 0.06]使用数字签名调整为[0.03, 0.03], 使用高斯机制调整为[0.004, 0.064], 篡改概率优化了 10.09%。

表 14 测定结果参数篡改概率和置信分数对比

Table 14 Comparison of tampering probability and confidence score of measurement result parameters (%)

p_i	篡改概率 PT			置信分数 SC		
	SHA3–256	Laplace	对比	SHA3–256	Laplace	对比
t_{Mx1}	30.83	30.00	–0.83	99.83	99.81	–0.02
t_{Inc1}	17.78	28.89	11.11	99.99	99.89	–0.10
t_{Wsh1}	20.00	5.00	–15.00	99.68	99.67	–0.01
t_{Mx2}	33.33	10.00	–23.33	98.95	98.50	–0.45
t_{Inc2}	38.33	5.00	–33.33	98.65	99.83	1.18
t_{Wsh2}	26.67	21.67	–5.00	98.53	99.83	1.30
t_{Mx3}	25.00	28.33	3.33	100.00	99.17	–0.83
t_{Inc3}	50.00	11.67	–38.33	100.00	99.67	–0.33
t_{Mx4}	21.00	21.50	0.50	99.83	99.55	–0.28
t_{Inc4}	48.89	28.89	–20.00	99.38	99.89	0.51
平均	31.18	19.09	–12.09	99.48	99.58	0.10

(注: 在拉普拉斯机制中取隐私预算 ε = 0.1.)

表 15 校准范围参数实验结果对比

Table 15 Comparison results of calibration range parameter

s_i	SHA3-256		Guass	
	数字签名	校准范围	高斯扰动	校准范围
[0,0.02]	110	[0.003,0.018]	[–0x4,0x5]	[0.006,0.024]
[0,0.04]	111	[0.003,0.037]	[–0x6,0xa]	[0.005,0.055]
[0,0.06]	111	[0.003,0.057]	[–0x8,0x9]	[0.004,0.064]
[0,0.08]	011	[0.001,0.077]	[–0x8,0xa]	[0.004,0.083]

(注: 在高斯机制中取隐私预算 $\varepsilon = 0.9$.)

表 16 校准范围参数篡改概率和校准度对比

Table 16 Comparison of tampering probability and calibration degree of the calibration range parameter (%)

s_i	篡改概率 PT			校准度 E		
	SHA3–256	Gaussian	对比	SHA3–256	Gaussian	对比
[0,0.02]	25.00	10.00	–15.00	75.00	90.00	15.00
[0,0.04]	15.00	3.25	–11.75	85.00	96.75	11.75
[0,0.06]	10.00	0.00	–10.00	90.00	100.00	10.00
[0,0.08]	5.00	1.38	–3.63	95.00	98.63	3.63
平均	13.75	3.66	–10.09	86.25	96.34	10.09

(注: 在高斯机制中取隐私预算 $\varepsilon = 0.9$.)

5.4.3 反馈信号参数对比结果

文献[17]设定蛋白质测定、插值混合以及两组模拟合成反应的随机检查点数量及误差限制阈值 $E_{\text{threshold}}$, 相同环境下, 应用本文所提的差分隐私机制在相同的生物芯片环境下所得累计误差率如表 17 所示, 对比文献[17]的误差限制阈值 $E_{\text{threshold}}$, 本文方案平均减少了 7.02%的累计误差率 e。

表 17 控制中心反馈信号误差率优化结果

Table 17 Optimization results of control center feedback signal error rate

生化协议	$E_{\text{threshold}}$(%)	检查点数量	误差率 e(%)	优化(%)
蛋白质测定	25.00	16	16.70	–8.30
	23.00	28	16.70	–6.30
	15.00	39	11.80	–3.20
插值混合	25.00	13	20.00	–5.00
	23.00	20	12.50	–10.50
	18.00	31	15.40	–2.60
合成 1	25.00	18	14.30	–10.70
	15.00	31	11.10	–3.90
合成 2	25.00	22	5.90	–19.10
	20.00	39	12.50	–7.50
	15.00	55	14.90	–0.10
平均				–7.02

(注: 在随机响应机制中取隐私预算参数 $p = 0.45$.)

6 总结

微流控平台在快速低成本生化分析方面具有巨大的潜力。然而, 使生物芯片自动化的网络物理系统反过来也让生物芯片遭受恶意攻击和知识产权盗窃。在网络威胁日益严重的时代, 安全是微流控技术大规模应用的主要障碍。上文详细讨论了差分隐私技术在生物芯片样品制备领域的应用场景, 并提出了参数安全发布算法。通过多轮仿真实验结果显示所提方案具有良好的隐私安全性和数据可用性。

参考文献

[1] Zhao Y, Chakrabarty K. Digital Microfluidic Logic Gates and Their Application to Built-in Self-Test of Lab-on-Chip[J]. *IEEE Transactions on Biomedical Circuits and Systems*, 2010, 4(4): 250-262.

[2] Zhong Z W, Li Z P, Chakrabarty K, et al. Micro-Electrode-Dot-Array Digital Microfluidic Biochips: Technology, Design Automation, and Test Techniques[J]. *IEEE Transactions on Biomedical Circuits and Systems*, 2019, 13(2): 292-313.

[3] Tang J, Ibrahim M, Chakrabarty K, et al. Toward Secure and Trustworthy Cyberphysical Microfluidic Biochips [J]. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2019, 38(4): 589-603.

[4] Dwork C, Roth A J F, science t i t c. The Algorithmic Foundations of Differential Privacy [J]. *The Algorithmic Foundations of Differential Privacy*, 2013.

[5] Li J D, Wang S J, Li K S M, et al. Watermarking for Paper-Based Digital Microfluidic Biochips[C]. *2020 IEEE International Test Conference in Asia*, 2020: 148-153.

[6] Shayan M, Bhattacharjee S, Tang J, et al. Bio-Protocol Watermarking on Digital Microfluidic Biochips[J]. *IEEE Transactions*

on Information Forensics and Security, 2019, 14(11): 2901-2915.

[7] Hsieh C W, Li Z P, Ho T Y. Piracy Prevention of Digital Microfluidic Biochips[C]. *2017 22nd Asia and South Pacific Design Automation Conference*, 2017: 512-517.

[8] Li J D, Wang S J, Li K S M, et al. Digital Rights Management for Paper-Based Microfluidic Biochips[C]. *2018 IEEE 27th Asian Test Symposium*, 2018: 179-184.

[9] Ali S S, Ibrahim M, Sinanoglu O, et al. Security Implications of Cyberphysical Digital Microfluidic Biochips[C]. *2015 33rd IEEE International Conference on Computer Design*, 2015: 483-486.

[10] He H, Hu H P. Field-Level Digital Microfluidic Biochips Trojan Detection Based on Hamming Distance[C]. *2020 IEEE 4th Information Technology*, *Networking*, *Electronic and Automation Control Conference*, 2020: 640-643.

[11] Shayan M, Bhattacharjee S, Song Y A, et al. Security Assessment of Microfluidic Immunoassays[C]. *The International Conference on Omni-Layer Intelligent Systems*, 2019: 217-222.

[12] Chen H L, Potluri S, Koushanfar F. BioChipWork: Reverse Engineering of Microfluidic Biochips[C]. *2017 IEEE International Conference on Computer Design*, 2017: 9-16.

[13] Bhattacharjee S, Tang J, Poddar S, et al. Bio-Chemical Assay Locking to Thwart Bio-IP Theft[J]. *ACM Transactions on Design Automation of Electronic Systems*, 2020, 25(1): 1-20.

[14] Ali S S, Ibrahim M, Sinanoglu O, et al. Microfluidic Encryption of On-Chip Biochemical Assays[C]. *2016 IEEE Biomedical Circuits and Systems Conference*, 2016: 152-155.

[15] Tang J, Member S, IEEE, et al. Secure Randomized Checkpointing for Digital Microfluidic Biochips [J]. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2017, PP(99): 1-1.

[16] [16] Yuan S L, Pi D C, Xu M. Trajectory Privacy Protection Method Based on Differential Privacy[J]. *Acta Electronica Sinica*, 2021, 49(7): 1266-1273.
(袁水莲, 皮德常, 胥萌. 基于差分隐私的轨迹隐私保护方法[J]. *电子学报*, 2021, 49(7): 1266-1273.)

[17] Zhao Y, Chakrabarty K. Integrated Control-Path Design and Error Recovery[M]. Design and Testing of Digital Microfluidic Biochips. New York, NY: Springer New York, 2012: 179-199.

[18] Pollack M G, Shenderov A D, Fair R B. Electrowetting-Based Actuation of Droplets for Integrated Microfluidics[J]. *Lab on a Chip*, 2002, 2(2): 96-101.

2002, 2(2): 96-101.

[19] Ali S S, Ibrahim M, Sinanoglu O, et al. Security Assessment of Cyberphysical Digital Microfluidic Biochips[J]. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2016, 13(3): 445-458.

[20] Alphonsus, H., C., et al. Digital Microfluidic Magnetic Separation for Particle-Based Immunoassays[J]. *Analytical Chemistry*, 2012, 84(20): 8805-8812.

[21] Dwork C. Differential privacy: A survey of results[C]. *The Theory and Applications of Models of Computation: 5th International Conference, TAMC 2008*, 2008: 1-19.

[22] Hu A T, Hu A Q, Hu Y, et al. Differentially Private Data Sharing and Publishing in Machine Learning: Techniques, Applications, and Challenges[J]. *Journal of Cyber Security*, 2022, 7(4): 1-16.
(胡奥婷, 胡爱群, 胡韵, 等. 机器学习中差分隐私的数据共享及发布: 技术、应用和挑战[J]. *信息安全学报*, 2022, 7(4): 1-16.)

[23] Bhattacharjee S, Banerjee A, Bhattacharya B B. Sample Preparation with Multiple Dilutions on Digital Microfluidic Biochips[J]. *IET Computers & Digital Techniques*, 2014, 8(1): 49-58.

[24] Luo Y, Chakrabarty K, Ho T Y. Error Recovery in Cyberphysical Digital Microfluidic Biochips[J]. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2013, 32(1): 59-72.

[25] Erlingsson Ú, Pihur V, Korolova A. RAPPOR: Randomized Aggregatable Privacy-Preserving Ordinal Response[C]. *The 2014 ACM SIGSAC Conference on Computer and Communications Security*, 2014: 1054-1067.

[26] Paul S, Mishra S. ARA: Aggregated RAPPOR and Analysis for Centralized Differential Privacy[J]. *SN Computer Science*, 2019, 1(1): 22.

[27] Fu Y, Yu Y H, Wu X P. Differential Privacy Protection Technology and Its Application in Big Data Environment[J]. *Journal on Communications*, 2019, 40(10): 157-168.
(付钰, 俞艺涵, 吴晓平. 大数据环境下差分隐私保护技术及应用[J]. *通信学报*, 2019, 40(10): 157-168.)

[28] Roy P, Banerjee A. A New Approach for Root-Causing Attacks on Digital Microfluidic Devices[C]. *2016 IEEE Asian Hardware-Oriented Security and Trust*, 2016: 1-6.

陈潇 于 2023 年在福州大学软件工程专业获得硕士学位。研究领域为生物芯片安全、差分隐私。研究兴趣包括: 机器学习隐私保护、数据安全。Email: 200327158@fzu.edu.cn

董晨 于 2011 年在武汉大学计算机应用技术专业获得博士学位。现任福州大学计算机与大数据学院硕士生导师。研究领域为芯片安全、智能计算。研究兴趣包括: 大数据、网络空间安全、人工智能、智能机器人。Email: dongchen@fzu.edu.cn